Differential Privacy:
Communications and Dissemination with the Technical and Stakeholder Community

Carol Hafford

Census National Advisory Committee

Implementing Differential Privacy for the 2020 Census Data Products Working Group

March 7, 2022

United States® Census Bureau

Discussant Topics

1. Learning from the Handbook for future communications

2. Key messages to multiple audiences

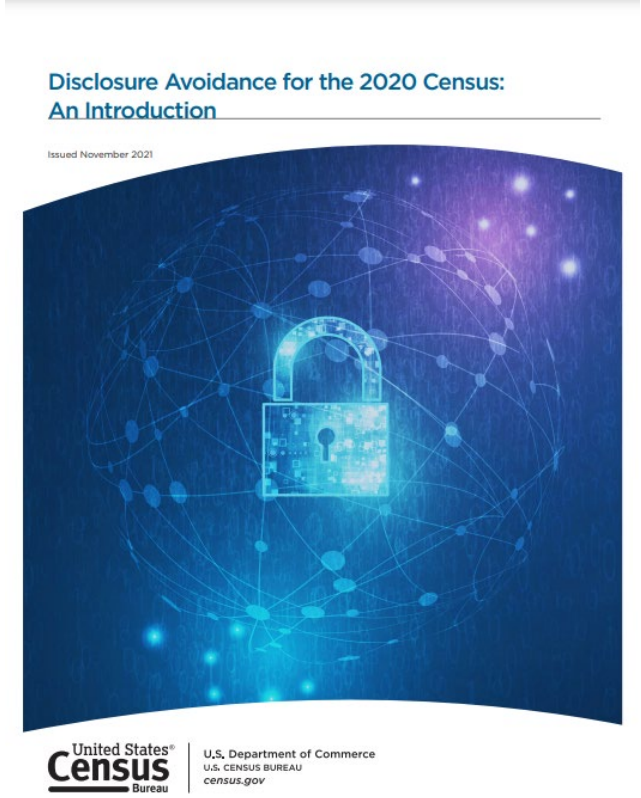
3. Increasing readability, engagement, and utility

4. *Discussion Questions:* Messaging and communications to technical and stakeholder audiences

United States® Census Bureau

The Handbook

1. Disclosure Avoidance for 2020 Census Redistricting Data: An Introduction
2. How Does the Disclosure Avoidance System Work for Redistricting Data?
3. Recommendations and Considerations When Using the Redistricting Data
4. Evaluating the 2020 Census Data
5. Frequently Asked Questions
6. Additional Resources
7. Glossary
8. Technical Appendix A: The Privacy-Loss Budget for 2020 Redistricting Data



https://www.census.gov/library/publications/2021/decennial/2020-census-disclosure-avoidance-handbook.html

United States® Census Bureau

What the Handbook does well … and can be replicated for future communications/products

- Sets the case for use of differential privacy, discussing the tradeoff between data confidentiality and data utility
- Identifies the advantages of using differential privacy as it offers quantifiable and provable confidentiality guarantees
- Places use of differential privacy in historical context, emphasizing the confluence of increased computing power, technological sophistication, and data availability, leading to the increased threat of reidentification attacks

United States® Census Bureau

Emphasize key messages

- The Bureau's Data Stewardship role
  - *Dual mandate to produce quality statistical information and protect the confidentiality of respondent data*
  - *Need to produce and deliver timely and accurate information to make informed decisions*
  - *Safeguard people's privacy and the confidentiality of their data*
- Limitations of differential privacy (e.g., impact on smaller geographies)
- What data users need to know before they start using statistics from the 2020 Census data files

United States® Census Bureau

Prioritize the message to identified audience

- **Establish the need for and rationale for differential privacy sooner**
  Why now?
  - *Reconstructing 100 percent of the 2010 Census records with full accuracy for 46 percent of the U.S. population is alarming.*
  - *Controlling the rate of inference an attacker may try to make about individuals is exactly the problem that the 2020 Census DAS was designed to address.*

- **Thematically organize the FAQs to address:**
  - Risks and known harms
  - Use of differential privacy as a protective measure
  - Types of data user engagement and concerns
  - Ability to compare data over time (i.e., 2010 and 2020 Census, 2020 Census and ACS)

- **Be intentional about the intended audiences**
  - "An Introduction" for whom?
  - Identify audience and appropriate level of proficiency for this and similar products with tiered messaging
    https://aircloak.com/explaining-differential-privacy
    - Easy
    - Intermediate
    - Professional

United States® Census Bureau

Increase readability, engagement, and utility

- Explain **key concepts** in **plain** *and* **statistical language** in the Glossary
  - Differential
  - Epsilon
  - Noise
  - Privacy loss budget
- Use **hyperlinked text** to allow the reader to access definitions in the Glossary
- Provide the Handbook in **on-line modular format**
- Present examples or use cases as **exhibits** (not as long passages in the text)
  Differential Privacy: A Primer for a Non-Technical Audience by Alexandra Wood, Micah Altman et al. (vanderbilt.edu)
  - Make them simple to understand while technically accurate, tailored to technical and stakeholder concerns
  - Use diverse names in examples to represent diverse racial/ethnic populations
- Tell visual stories about the data and differential privacy
  - Use **infographics** to clearly and succinctly communicate data concepts, decision rules, and transformation
  - Create **interactive data visualization tools** to increase user engagement at all levels of proficiency

  Examples from the Handbook:
  - To explain the tradeoff in the Privacy Loss Budget and how it is set (e.g., a dial that tunes the amount of noise that is added to the data)
  - To depict the invariant statistics for the 2020 Census redistricting data are and where noise has been introduced
  - To convey *How Does the Top Down Algorithm (TDA) Work*?
- Lay out **advantages** and **limitations** (side by side) to enhance transparency and foster public discussion

United States® Census Bureau

? *NAC Discussion:* Messaging and communications to technical and stakeholder audiences

To date, the Bureau has met with multiple stakeholders, including the scientific and statistical community, federal partners, tribal leaders, state and local officials, civil rights groups, academic researchers, etc.

- What concerns did they raise?
  - Disparate impact on level of geography?
  - Disparate impact on racial/ethnic groups?
  - Other impact?
- How did the Bureau respond to these concerns?
- What feedback has the Bureau received on data utility and validity from stakeholders?
  - How will that feedback be used to create new products that explain the need for differential privacy, its advantages, and limitations?

Moving forward, what is the Bureau's technical and stakeholder engagement strategy?

- Does it address outreach, consultations, involvement, collaboration across technical and stakeholder groups?
  - Building on the stakeholder convenings to date, who else will the Bureau talk to? And when?
- Has there been consideration of tailored products the Bureau might create to address concerns of varied data user and stakeholder groups?
- What stakeholder advice and lessons learned from the original round of data user engagement can inform how differential privacy will be messaged?
- How will technical and stakeholder groups be involved in message design and communications?

United States® Census Bureau

COMMITTEE DISCUSSION

For audio to this conference, please dial 1-888-566-6192 (Access Code: 4881978#)

United States® Census Bureau